Journal of Multivariate Analysis **75**, 143–162 (2000)
doi:10.1006/jmva.1999.1890, available online at http://www.idealibrary.com on IDEAL®

Statistical Aspects of Perpetuities

Rudolf Grübel

Universität Hannover, Hannover, Germany
E-mail: rgrubel@stochastik.uni-hannover.de

and

Susan M. Pitts

University of Cambridge, Cambridge, United Kingdom
E-mail: s.pitts@statslab.cam.ac.uk

For a distribution μ on the unit interval we define the associated perpetuity $\Psi(\mu)$ as the distribution of $1+X_1+X_1X_2+X_1X_2X_3+\cdots$, where $(X_n)_{n\in\mathbb{N}}$ is a sequence of independent random variables with distribution μ. Such quantities arise in insurance mathematics and in many other areas. We prove the differentiability of the *perpetuity functional* Ψ with respect to integral and supremum norms. These results are then used to investigate the statistical properties of empirical perpetuities, including the behaviour of bootstrap confidence regions. © 2000 Academic Press

AMS 1980 subject classifications: 62G05, 62G09, 62P05.
Key words and phrases: perpetual annuity, empirical perpetuities, asymptotic normality, bootstrap.

1. INTRODUCTION

Let μ be a probability distribution on the unit interval $[0, 1]$; to avoid trivialities we will assume throughout the paper that μ is not concentrated on the single value 1. It is then easy to see that for any sequence $(X_n)_{n\in\mathbb{N}}$ of independent random variables with distribution μ the cumulative products $(X_1\cdot\cdots\cdot X_n)_{n\in\mathbb{N}}$ of the sequence sum to a finite value with probability one; we call the distribution of $Y:=1+\sum_{n=1}^{\infty}\prod_{m=1}^{n}X_m$ the *perpetuity* associated with μ and denote it by $\Psi(\mu)$.

Perpetuities arise in a variety of problems in pure and applied mathematics. The name points to a financial context: Y can be regarded as the present value of a perpetual periodic payment of one monetary unit if X_n

0047-259X/00 $35.00

denotes the value at the beginning of time period n of one unit paid out at the beginning of time period $n+1$; see Dufresne (1990) for the risk and insurance theory context and a review of the relevant literature. Perpetuities also appear in the analysis of a random algorithm where μ is the uniform distribution on (0, 1) or (1/2, 1); see Grübel and Rösler (1996) and Grübel (1998). The perpetuity associated with the uniform distribution on the unit interval also arises in number theory; see de Bruijn (1951). Further, $\Psi(\mu)$ is the stationary distribution of a system $(Y_n)_{n\in\mathbb{N}}$ with dynamic evolution $Y_{n+1} = 1 + X_{n+1} Y_n$. Random affine maps and stochastic difference equations provide further areas where perpetuities appear naturally; see Vervaat (1979), Goldie and Grübel (1996) and the references given there for details and more applications. Of course, in many applications generalizations of the above setup are needed, with, e.g., μ not concentrated on the unit interval, or matrix-valued X-variables.

Interest in the literature focuses on convergence issues and tail probabilities. Despite their practical importance, statistical aspects of perpetuities seem to have received little attention so far, the most notable exception we are aware of being the paper by Aebi *et al.* (1994). In the present paper we show that a moderately abstract viewpoint is useful in this context. We regard Ψ as a nonlinear operator, the *perpetuity functional*, mapping probability distributions to probability distributions. As a first application and on a heuristic level, this functional approach provides a natural estimator for $\Psi(\mu)$, given a sample $X_1, ..., X_n$ from μ. With $\hat{\mu}_n$ denoting the empirical distribution associated with the sample, i.e., the probability measure that assigns mass $1/n$ to each of $X_1, ..., X_n$, the "plug-in principle" leads us to estimate $\Psi(\mu)$ by $\Psi(\hat{\mu}_n)$, the *empirical perpetuity*. This is, essentially, the "bootstrap estimator" investigated by Aebi *et al.* (1994).

A suitable continuity property of the perpetuity functional together with the known consistency of $\hat{\mu}_n$ as an estimator of μ implies the consistency of empirical perpetuities. Further, on the next level of detail, a suitable differentiability property of the functional together with the known asymptotic distributional behaviour of empirical distributions can be used to prove asymptotic normality of $\Psi(\hat{\mu}_n)$ (the von Mises or delta method). The same differentiability property can be used to show that "the bootstrap works."

The precise results are given in Section 2. Section 3 contains the proofs. In previous work we have used this functional approach in a variety of situations; see Grübel and Pitts (1993) for renewal theory, Pitts (1994a) for the context of queueing theory, and Pitts (1994b) for compound distributions. In these applications convolution series appear directly or indirectly, and Banach algebra theory is an important tool. The situation considered here is of a completely different type, with the analysis of fixed point equations playing a central role.

2. RESULTS

Let $(X_i)_{i\in\mathbb{N}}$, μ, $\hat{\mu}_n$ be as in the Introduction, and let $\hat{G}_n$ be the distribution function associated with the empirical perpetuity $\Psi(\hat{\mu}_n)$. Similarly, G denotes the distribution function associated with the unknown perpetuity $\Psi(\mu)$. Our interest concentrates on two statements: first, asymptotic normality of empirical perpetuities, and second, asymptotic validity of bootstrap confidence regions for the unknown perpetuity.

Formally, the first statement means that

$$\sqrt{n}\,(\hat{G}_n - G) \to_d Z \qquad \text{as} \qquad n \to \infty, \tag{AN}$$

with some Gaussian process Z. Here $\to_d$ denotes convergence in distribution; in particular, (AN) refers to a topology on a suitable space $\mathbb{F}$ of functions, and our results give conditions for the validity of (AN) in two different classes of normed function spaces. In both cases we regard Ψ as an operator from $D[0, 1]$, the space of cadlag functions f: $[0, 1] \to \mathbb{R}$, to $\mathbb{F}$, and compute its derivative Ψ'_μ at μ. Let F be the distribution function associated with μ and let $B = (B_t)_{0 \leqslant t \leqslant 1}$ be the Brownian bridge. The limit process Z in (AN) is the image of the time-changed Brownian bridge $B \circ F = (B_{F(t)})_{0 \leqslant t \leqslant 1}$ under the bounded linear map Ψ'_μ.

A consequence of (AN) of particular interest is the convergence of

$$R_n(x) = P(\sqrt{n}\,\|\hat{G}_n - G\| \leqslant x)$$

to the analogous quantity associated with the limit process,

$$R(x) = P(\|Z\| \leqslant x),$$

for all continuity points x of the latter. Quantiles of R could be used to set up asymptotic confidence regions for G. This would, however, involve the calculation of the distribution of the supremum of the Gaussian process Z, whose second order structure depends on the underlying unknown μ in a complicated way—with $\mathbb{F}'$ and $D[0, 1]'$ the topological duals of $\mathbb{F}$ and $D[0, 1]$, respectively, and $\Psi'^{\star}_\mu$: $\mathbb{F}' \to D[0, 1]'$ the adjoint of Ψ'_μ, the covariance function of Z would be given by

$$\operatorname{cov}(\langle f, Z\rangle, \langle g, Z\rangle) = E(\langle \Psi'^{\star}_\mu(f), B \circ F\rangle \cdot \langle \Psi'^{\star}_\mu(g), B \circ F\rangle)$$

for all $f, g \in \mathbb{F}'$. The bootstrap resolves this dilemma, using an estimate for R_n which is in turn obtained via the plug-in principle as the R_n-quantity associated with the empirical distribution. For a formally correct definition let $\mu_n(x_1, ..., x_n)$ be the probability measure that assigns mass $1/n$ to each of the values $x_1, ..., x_n$, i.e., with δ_x denoting unit mass at x,

$$\mu_n(x_1, ..., x_n) = \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i}.$$

With this notation we can regard the empirical distribution as the random probability measure $\hat{\mu}_n = \mu_n(X_1, ..., X_n)$. Then with $I_n := \{1, ..., n\}^n$,

$$\hat{R}_n(z) = n^{-n} \sum_{(i_1, ..., i_n) \in I_n} 1_{[0, z]}(\sqrt{n}\, \| \Psi(\mu_n(X_{i_1}, ..., X_{i_n})) - \Psi(\hat{\mu}_n) \|).$$

A numerical approximation for this quantity can be calculated by the usual Monte Carlo method. Note that this differs from Aebi *et al.* (1994), who investigated the variability of the Monte Carlo approximation to $\hat{G}_n$. We can now state our second aim, which is to establish that the bootstrap works. By this we mean that

$$\hat{R}_n \to R \qquad \text{in probability as} \quad n \to \infty. \tag{BW}$$

This implies that the quantiles associated with $\hat{R}_n$ (which can be calculated from the data to any desired degree of precision) can serve as a substitute for the quantiles associated with R_n if n is large, as both distribution functions have the same limit R.

If $q_{n, \alpha}$ is the (estimated) α-quantile of the distribution of the pivot, then the corresponding confidence region for G consists of all distribution functions G' with $\| \cdot \|$-distance from $\hat{G}_n$ less than or equal to $q_{n, \alpha}/\sqrt{n}$. Our first result uses a weighted integral norm and does not impose any additional restrictions on the underlying distribution μ. For this result we consider the quantity $\sqrt{n}\,(\hat{G}_n - G)$ as an element of the space

$$L_{1, \sigma} := \left\{ f\colon [1, \infty) \to \mathbb{R}\colon \int_1^\infty e^{\sigma x} \, |f(x)| \, dx < \infty \right\},$$

where σ depends on μ. With the usual identification of functions that differ only on a Lebesgue null set, $L_{1, \sigma}$ becomes a Banach space with the norm

$$\|f\|_{1, \sigma} := \int_1^\infty e^{\sigma x} \, |f(x)| \, dx.$$

Let $m_1(y) = \int x\mu(dx)$ be the first moment associated with μ.

THEOREM 1. *If* $\sigma < -\log m_1(\mu)$ *then* (AN) *and* (BW) *hold with respect to the* $\| \cdot \|_{1, \sigma}$*-norm.*

Confidence regions based on the $\|\cdot\|_{1,\sigma}$-pivot are not easily visualised. Pivots based on the supremum norm lead to confidence bands for the distribution function of the perpetuity, which can be displayed easily and could also be used to set up confidence intervals for e.g. quantiles of the perpetuity. We consider supremum norms with a continuous and increasing weight function $\rho: [1, \infty) \to (0, \infty)$, which we further assume to satisfy

$$\sup_{x \geqslant 0} \frac{\rho(1+x)}{\rho(1 \vee x)} < \infty. \tag{1}$$

Let D_0 be the set of all functions $f: [1, \infty] \to \mathbb{R}$ which are right continuous, have left-hand limits and satisfy $f(\infty) = f(\infty -) = 0$. Let D_ρ be the set of all $f \in D_0$ with

$$\|f\|_{\infty,\rho} := \sup_{x \geqslant 1} \rho(x)|f(x)| < \infty.$$

With its usual linear structure $(D_\rho, \|\cdot\|_{\infty,\rho})$ is a Banach space.

THEOREM 2. *If ρ is such that $\rho(x) = O(e^{\sigma x})$ for some $\sigma < -\log m_1(\mu)$ as $x \to \infty$ and if*

$$C(\mu, \rho) := \sup_{x \geqslant 0} \rho(1+x) \int_{(0,\, 1 \wedge x]} \frac{1}{\rho(x/y)} \mu(dy) < 1,$$

then (AN) *and* (BW) *hold with respect to the $\|\cdot\|_{\infty,\rho}$-topology.*

The condition $C(\mu, \rho) < 1$ will be discussed in Section 3.5, where we show that it is satisfied with $\rho(x) = \exp(\sigma x)$ for $\sigma < \log 2 \approx 0.693$ if μ has a decreasing density.

3. PROOFS

In Section 3.1 we introduce some notation and comment on general aspects of the proofs. In Sections 3.2 and 3.3 we prove the differentiability of Ψ with respect to the integral norm and the supremum norm respectively. Section 3.4 explains how these differentiability properties are connected to (AN) and (BW). In the final subsection we discuss the condition $C(\mu, \rho) < 1$ in Theorem 2.

3.1. For any pair μ, ν of finite signed measures, with μ concentrated on $[0, 1]$ and ν concentrated on $[1, \infty)$, let $T(\mu, \nu)$ be the finite signed measure concentrated on $[1, \infty)$ defined by

$$T(\mu, \nu)((x, \infty)) := \int_{(0,1]} \nu\left(\left(\frac{x-1}{y}, \infty\right)\right) \mu(dy) \qquad \text{for all} \quad x \geqslant 1, \tag{2}$$

and $T(\mu, \nu)(\{1\}) := \mu(\{0\})\, \nu([1, \infty))$. Obviously, this defines a bilinear map. A signed measure ν on $[1, \infty)$ of known total mass can be characterized by its *tail function* f: $[1, \infty) \to \mathbb{R}$, $f(x) := \nu((x, \infty))$. For fixed μ we could interpret $T(\mu, \cdot)$ as an operator mapping tail functions to tail functions. If ν has total mass 0 then

$$T(\mu, \nu)((x, \infty)) = \int_{(0, 1 \wedge (x-1)]} f\left(\frac{x-1}{y}\right) \mu(dy) \qquad \text{for all} \quad x \geqslant 1, \tag{3}$$

and the total mass of $T(\mu, \nu)$ is again 0. Using Fubini's theorem, we can rewrite (2) as

$$T(\mu, \nu)((x, \infty)) = \int_{[1, \infty)} \mu\left(\left(\frac{x-1}{y}, 1\right]\right) \nu(dy) \qquad \text{for all} \quad x \geqslant 1, \tag{4}$$

and if f denotes the tail function of a signed measure on $[0, 1]$ with total mass 0 then, similar to the transition from (2) to (3),

$$T(\mu, \nu)((x, \infty)) = \int_{[(x-1) \vee 1, \infty)} f\left(\frac{x-1}{y}\right) \nu(dy) \qquad \text{for all} \quad x \geqslant 1. \tag{5}$$

Equations (3) and (5) show that T can be extended to operate on pairs of arguments consisting of a finite signed measure and a cadlag function. Below we will often identify finite signed measures of known total mass with their tail function.

If μ and ν are the distributions of independent random variables X and Y respectively then $T(\mu, \nu)$ is the distribution of $1 + X \cdot Y$. This implies the following important relationship between Ψ and T,

$$T(\mu, \Psi(\mu)) = \Psi(\mu), \tag{6}$$

whenever $\Psi(\mu)$ is defined. Equation (6) shows that $\Psi(\mu)$ can be regarded as the fixed point of the linear map $\nu \mapsto T(\mu, \nu)$. Using the bilinearity of T we obtain for any μ_1, μ_2

$$\Psi(\mu_1) - \Psi(\mu_2) = T(\mu_1 - \mu_2, \Psi(\mu_1)) + T(\mu_2, \Psi(\mu_1) - \Psi(\mu_2)). \tag{7}$$

If we have a curve $\{\mu_\varepsilon\}_{0\leqslant\varepsilon\leqslant 1}$ in the space of probability measures we similarly obtain

$$\frac{1}{\varepsilon}(\Psi(\mu_\varepsilon)-\Psi(\mu_0))=T\left(\frac{1}{\varepsilon}(\mu_\varepsilon-\mu_0),\ \Psi(\mu_\varepsilon)\right)+T\left(\mu_0,\frac{1}{\varepsilon}(\Psi(\mu_\varepsilon)-\Psi(\mu_0))\right). \tag{8}$$

This exhibits the difference quotient $\varepsilon^{-1}(\Psi(\mu_\varepsilon)-\Psi(\mu_0))$ as a fixed point of the affine operator $v\mapsto T(\varepsilon^{-1}(\mu_\varepsilon-\mu_0),\ \Psi(\mu_\varepsilon))+T(\mu_0,v)$. This will be important in proving differentiability of Ψ, by which we mean that the convergence of $\varepsilon^{-1}(\mu_\varepsilon-\mu_0)$ as $\varepsilon\to 0$ to some object g implies the convergence of the associated difference quotients to some $h=\Psi'_{\mu_0}(g)$, the linear operator Ψ'_{μ_0} being the derivative of Ψ at μ_0. The topologies will depend on the "point" μ_0 where we want to differentiate Ψ.

Before we begin with the technical details of the proofs one more general comment is in order. Equation (6) also implies

$$\Phi(\mu,\Psi(\mu))=0,\qquad \text{with}\quad \Phi(\mu,v):=v-T(\mu,v).$$

Written in this form, proving differentiability of Ψ looks like a straightforward case for an implicit function theorem. However, the statistical applications dictate a weak topology, where the arguments of T could leave the space of signed measures (Brownian motion paths are not of bounded variation), and it is not clear how to define T if neither argument is the distribution function or, equivalently, the tail function of a signed measure. In our more direct approach we avoid this problem. As a rough guideline to the proofs it is perhaps worth mentioning in this context that the implicit function theorem would require the invertibility of the derivative Φ_2 of Φ with respect to the second argument. The obvious candidate is $\Phi_2=\mathrm{Id}-T(\mu,\cdot)$, which leads us to consider the Neumann series associated with the linear operator $T(\mu,\cdot)$.

3.2. For a probability measure μ on $[0,1]$ with $\mu(\{1\})\neq 1$ and f in $L_{1,\sigma}$, let

$$(U_\mu f)(x)=\int_{(0,\,1\,\wedge\,(x-1)]} f\left(\frac{x-1}{y}\right)\mu(dy),\qquad x\geqslant 1.$$

If v has tail function f and total mass 0 then, from (3),

$$(U_\mu f)(x)=T(\mu,v)((x,\infty)). \tag{9}$$

On the space $D[0,1]$ of cadlag functions $f\colon[0,1]\to\mathbb{R}$ we consider the norm $\|f\|_\infty:=\sup_{0\leqslant x\leqslant 1}|f(x)|$. We have the following basic inequalities.

LEMMA 3. (i) *For all* $f \in L_{1,\sigma}$,

$$\|U_\mu f\|_{1,\sigma} \leqslant e^\sigma m_1(\mu)\, \|f\|_{1,\sigma}.$$

(ii) *For all* $f \in D[0,1]$ *and all probability measures* ν *concentrated on* $[1,\infty)$,

$$\|T(f,\nu)\|_{1,\sigma} \leqslant e^\sigma\, \|f\|_\infty \left(\|\nu\|_{1,\sigma} + \frac{e^\sigma}{\sigma}\right).$$

Proof. The proof of (i) follows from

$$\begin{aligned}
\int_1^\infty e^{\sigma x} |U_\mu f(x)|\, dx &\leqslant \int_1^\infty e^{\sigma x} \int_{(0,\, 1 \wedge (x-1)]} \left| f\left(\frac{x-1}{y}\right)\right| \mu(dy)\, dx \\
&= \int_{(0,1]} \int_{y+1}^\infty e^{\sigma x} \left| f\left(\frac{x-1}{y}\right)\right| dx\, \mu(dy) \\
&= e^\sigma \int_{(0,1]} y \int_1^\infty e^{\sigma u y}\, |f(u)|\, du\, \mu(dy) \\
&\leqslant e^\sigma m_1(\mu)\, \|f\|_{1,\sigma};
\end{aligned}$$

for the proof of (ii) we use

$$\begin{aligned}
\|T(f,\nu)\|_{1,\sigma} &\leqslant \int_1^\infty e^{\sigma x} \int_{[(x-1) \vee 1,\, \infty)} \left| f\left(\frac{x-1}{y}\right)\right| \nu(dy)\, dx \\
&= \int_{[1,\infty)} \int_1^{y+1} e^{\sigma x} \left| f\left(\frac{x-1}{y}\right)\right| dx\, \nu(dy) \\
&= \int_{[1,\infty)} y \int_0^1 e^{\sigma(uy+1)}\, |f(u)|\, du\, \nu(dy) \\
&\leqslant \|f\|_\infty \int_{[1,\infty)} \int_1^{y+1} e^{\sigma z}\, dz\, \nu(dy) \\
&= \|f\|_\infty \int_1^\infty e^{\sigma z} \int_{[(z-1)\vee 1,\, \infty)} \nu(dy)\, dz \\
&= \|f\|_\infty \left(\int_1^2 e^{\sigma z} \nu([1,\infty))\, dz + \int_1^\infty e^{\sigma(z+1)} \nu([z,\infty))\, dz \right) \\
&\leqslant \|f\|_\infty\, e^\sigma \left(\frac{e^\sigma}{\sigma} + \|\nu\|_{1,\sigma}\right). \quad \blacksquare
\end{aligned}$$

Note that (i) implies that U_μ is a bounded linear map from $L_{1,\sigma}$ to $L_{1,\sigma}$. Further, if $e^{\sigma} m_1(\mu) < 1$ then U_μ is a contraction, and the partial sums $\sum_{k=0}^{n} U_\mu^k f$, $n \in \mathbb{N}$, constitute a Cauchy sequence in $(L_{1,\sigma}, \|\cdot\|_{1,\sigma})$ (here U_μ^0 is understood to be the identity on $L_{1,\sigma}$). We can therefore define another linear operator $V_\mu : L_{1,\sigma} \to L_{1,\sigma}$ by $V_\mu := \sum_{k=0}^{\infty} U_\mu^k$. Simple standard arguments from functional analysis suffice to prove the following properties of the Neumann series V_μ (see Heuser, 1982, Section 8).

Lemma 4. (i) *V_μ is continuous.*

(ii) *If $f, g \in L_{1,\sigma}$, are such that $f = g + U_\mu f$, then $f = V_\mu g$.*

For the following lemmas we consider a family $\{\mu_\varepsilon\}_{0 \leqslant \varepsilon \leqslant 1}$ of probability measures on $[0, 1]$. We assume that $\mu_0(\{1\}) \neq 1$ and fix some σ, $0 < \sigma < -\log m_1(\mu_0)$. We begin our local analysis of the functional Ψ by proving local boundedness.

Lemma 5. *If*

$$\lim_{\varepsilon \to 0} \sup_{0 \leqslant x \leqslant 1} |\mu_\varepsilon((x, 1]) - \mu_0((x, 1])| = 0,$$

then

$$\limsup_{\varepsilon \to 0} \|\Psi(\mu_\varepsilon)\|_{1,\sigma} < \infty.$$

Proof. Choose $\theta_0 \in (\sigma, -\log m_1(\mu_0))$ and $\eta \in (m_1(\mu_0), 1)$ such that $e^{\theta_0}\eta < 1$. Now choose $\kappa > 0$ such that

$$e^{\theta}(1 + \kappa\theta\eta) \leqslant 1 + \kappa\theta \qquad \text{for} \quad 0 \leqslant \theta \leqslant \theta_0. \tag{10}$$

Let $\varepsilon_0 > 0$ be such that $m_1(\mu_\varepsilon) \leqslant \eta$ for all $\varepsilon \leqslant \varepsilon_0$ (note that the assumptions on the support of the measures imply convergence of the first moments). For a probability measure μ let M_μ, $M_\mu(\theta) := \int e^{\theta x} \mu(dx)$ be the associated moment generating function; if $\mu = \mathscr{L}(X)$ we simply write M_X. We claim that

$$M_{\Psi(\mu_\varepsilon)}(\theta_0) \leqslant 1 + \kappa\theta_0 \qquad \text{for all} \quad \varepsilon \leqslant \varepsilon_0. \tag{11}$$

To prove this for a given $\varepsilon \leqslant \varepsilon_0$ we note that $\Psi(\mu_\varepsilon)$ is the distribution of the random variable $Y = 1 + \sum_{k=1}^{\infty} \prod_{i=1}^{k} X_i$ with $\{X_i\}$ independent, $\mathscr{L}(X_i) = \mu_\varepsilon$. This variable is the monotone limit of the variables $Y_n := 1 + \sum_{k=1}^{n} \prod_{i=1}^{k} X_i$ as $n \to \infty$, so (11) follows if we can show that

$$M_{Y_n}(\theta) \leqslant 1 + \kappa\theta \qquad \text{for all} \quad \theta \in [0, \theta_0], \qquad n \in \mathbb{N}. \tag{12}$$

For $n=0$ this is obvious from (10) since $Y_0 \equiv 1$. We have

$$Y_{n+1} = 1 + X_1 \cdot \left(1 + \sum_{k=1}^{n} \prod_{i=1}^{k} X_{i+1}\right),$$

which implies $\mathscr{L}(Y_{n+1}) = \mathscr{L}(1 + X_{n+1} Y_n)$. Therefore, if (12) holds for some n, then

$$\begin{aligned} M_{Y_{n+1}}(\theta) &= E \exp(\theta(1 + X_{n+1} Y_n)) \\ &= e^{\theta} \int_{[0,1]} M_{Y_n}(x\theta)\, \mu_\varepsilon(dx) \\ &\leqslant e^{\theta} \int_{[0,1]} (1 + \kappa x \theta)\, \mu_\varepsilon(dx) \\ &= e^{\theta}(1 + \kappa\theta m_1(\mu_\varepsilon)) \\ &\leqslant 1 + \kappa\theta. \end{aligned}$$

This completes the inductive proof of (12), and hence (11). Using

$$\begin{aligned} \|\Psi(\mu_\varepsilon)\|_{1,\sigma} &= \int_1^\infty e^{\sigma x} \int_{(x,\infty)} \Psi(\mu_\varepsilon)(dy)\, dx \\ &= \int_{(1,\infty)} \int_1^y e^{\sigma x}\, dx\, \Psi(\mu_\varepsilon)(dy) \\ &\leqslant \frac{1}{\sigma} M_{\Psi(\mu_\varepsilon)}(\sigma) \\ &\leqslant \frac{1}{\sigma} M_{\Psi(\mu_\varepsilon)}(\theta_0), \end{aligned}$$

we see that this implies the statement of the lemma. ∎

Next we prove continuity. As we deal with the limit $\varepsilon \to 0$, we may assume in the proofs below that ε is small enough for the respective quantities to be defined.

LEMMA 6. *If*

$$\lim_{\varepsilon \to 0} \sup_{0 \leqslant x \leqslant 1} |\mu_\varepsilon((x,1]) - \mu_0((x,1])| = 0,$$

then

$$\lim_{\varepsilon \to 0} \|\Psi(\mu_\varepsilon) - \Psi(\mu_0)\|_{1,\sigma} = 0.$$

Proof. From (7),

$$\Psi(\mu_\varepsilon) - \Psi(\mu_0) = T(\mu_\varepsilon - \mu_0, \Psi(\mu_\varepsilon)) + T(\mu_0, \Psi(\mu_\varepsilon) - \Psi(\mu_0)),$$

and, by (9), this can be written

$$\Psi(\mu_\varepsilon) - \Psi(\mu_0) = g + U_{\mu_0}(\Psi(\mu_\varepsilon) - \Psi(\mu_0)),$$

where $g = T(\mu_\varepsilon - \mu_0, \Psi(\mu_\varepsilon))$. Applying Lemma 4 (ii), we find

$$\Psi(\mu_\varepsilon) - \Psi(\mu_0) = V_{\mu_0} T(\mu_\varepsilon - \mu_0, \Psi(\mu_\varepsilon)).$$

Then, since V_{μ_0} is bounded (by Lemma 4 (i)), we have

$$\begin{aligned}\|\Psi(\mu_\varepsilon) - \Psi(\mu_0)\|_{1,\sigma} &\leqslant \|V_{\mu_0}\| \; \|T(\mu_\varepsilon - \mu_0, \Psi(\mu_\varepsilon))\|_{1,\sigma} \\ &\leqslant \|V_{\mu_0}\| \; e^\sigma \; \|\mu_\varepsilon - \mu_0\|_\infty \left(\|\Psi(\mu_\varepsilon)\|_{1,\sigma} + \frac{e^\sigma}{\sigma}\right),\end{aligned}$$

by Lemma 3 (ii). The right-hand side tends to zero as ε tends to zero by the assumptions of this lemma and by Lemma 5. ∎

For the final step we need a suitable form of continuity of T with respect to its second argument; note that the inequality in Lemma 3 (ii) is not sufficient for this purpose.

LEMMA 7. *If $\{\nu_\varepsilon\}_{0 \leqslant \varepsilon \leqslant 1}$ is a family of probability measures in $L_{1,\sigma}$ with*

$$\lim_{\varepsilon \to 0} \|\nu_\varepsilon - \nu_0\|_{1,\sigma} = 0,$$

then, for any $f \in D[0, 1]$ with $f(1) = 0$,

$$\lim_{\varepsilon \to 0} \|T(f, \nu_\varepsilon) - T(f, \nu_0)\|_{1,\sigma} = 0.$$

Proof. We write 1_A for the indicator function of the set A. For $0 \leqslant a < b \leqslant 1$ we obtain

$$\begin{aligned}&\int_1^\infty e^{\sigma x} \left| \int_{[1,\infty)} 1_{[a,b)}\left(\frac{x-1}{y}\right) \nu_\varepsilon(dy) - \int_{[1,\infty)} 1_{[a,b)}\left(\frac{x-1}{y}\right) \nu_0(dy)\right| dx \\ &\quad\leqslant \int_{1+b}^\infty e^{\sigma x} \left| \nu_\varepsilon\left(\left(\frac{x-1}{b}, \infty\right)\right) - \nu_0\left(\left(\frac{x-1}{b}, \infty\right)\right)\right| dx \\ &\qquad + \int_{1+a}^\infty e^{\sigma x} \left| \nu_\varepsilon\left(\left(\frac{x-1}{a}, \infty\right)\right) - \nu_0\left(\left(\frac{x-1}{a}, \infty\right)\right)\right| dx \\ &\quad\leqslant (a+b)\, e^\sigma \, \|\nu_\varepsilon - \nu_0\|_{1,\sigma}.\end{aligned}$$

This shows that the statement of the lemma holds for $f=1_{[a,b)}$, from which it follows easily that it also holds for finite linear combinations of such indicator functions. Now let f be an arbitrary element of $D[0,1]$ with $f(1)=0$ and let

$$K := e^{\sigma} \sup_{0\leqslant\varepsilon\leqslant 1} \|v_\varepsilon\|_{1,\sigma} + \sigma^{-1}e^{2\sigma}.$$

For any given $\delta>0$ we can find a function g that can be written as a finite linear combination of indicator functions of intervals $[a,b)$, $0\leqslant a<b\leqslant 1$, and satisfies $\|f-g\|_\infty<\delta/(3K)$. For g we can find an $\varepsilon_0>0$ such that for all $\varepsilon\in(0,\varepsilon_0)$,

$$\|T(g,v_\varepsilon)-T(g,v_0)\|_{1,\sigma}\leqslant\delta/3.$$

Using this, the bound

$$\begin{aligned}&\|T(f,v_\varepsilon)-T(f,v_0)\|_{1,\sigma}\\&\qquad\leqslant\|T(f-g,v_\varepsilon)\|_{1,\sigma}+\|T(g,v_\varepsilon)-T(g,v_0)\|_{1,\sigma}+\|T(g-f,v_0)\|_{1,\sigma},\end{aligned}$$

and Lemma 3 (ii) we obtain

$$\|T(f,v_\varepsilon)-T(f,v_0)\|_{1,\sigma}\leqslant\delta$$

for all $\varepsilon\leqslant\varepsilon_0$. ∎

We finally arrive at the main technical result of this subsection, the differentiability of the perpetuity functional.

PROPOSITION 8. *If*

$$\lim_{\varepsilon\to 0}\sup_{0\leqslant x\leqslant 1}\left|\frac{1}{\varepsilon}(\mu_\varepsilon-\mu_0)((x,1])-g(x)\right|=0$$

for some $g\in D[0,1]$, *then*

$$\frac{1}{\varepsilon}(\Psi(\mu_\varepsilon)-\Psi(\mu_0))\to V_{\mu_0}T(g,\Psi(\mu_0)) \qquad \textit{in} \quad L_{1,\sigma}.$$

Proof. With our basic curve $\varepsilon\to\mu_\varepsilon$ we associate the following functions f_ε and h_ε,

$$\begin{aligned}f_\varepsilon(x) &:= \frac{1}{\varepsilon}(\Psi(\mu_\varepsilon)-\Psi(\mu_0))((x,\infty)),\\ h_\varepsilon(x) &:= T\left(\frac{1}{\varepsilon}(\mu_\varepsilon-\mu_0),\Psi(\mu_\varepsilon)\right)((x,\infty)), \qquad \text{for all} \quad x\geqslant 1.\end{aligned}$$

From (8) and (9) we have $f_\varepsilon = h_\varepsilon + U_{\mu_0} f_\varepsilon$ so that $f_\varepsilon = V_{\mu_0} h_\varepsilon$ by Lemma 4(ii). The decomposition

$$h_\varepsilon - T(g, \Psi(\mu_0)) = T\left(\frac{1}{\varepsilon}(\mu_\varepsilon - \mu_0), \Psi(\mu_\varepsilon)\right) - T(g, \Psi(\mu_\varepsilon)) \\ + T(g, \Psi(\mu_\varepsilon)) - T(g, \Psi(\mu_0)),$$

together with Lemma 3 (ii) and Lemma 7 shows that $T(g, \Psi(\mu_0))$ is the limit of h_ε as $\varepsilon \to 0$. The statement of the proposition now follows on using the continuity of V_{μ_0}. ∎

We may condense this proposition into a single formula,

$$\Psi'_{\mu_0} = V_{\mu_0} T(\cdot, \Psi(\mu_0)).$$

3.3. Our aim in this subsection is to prove the differentiability of Ψ, now regarded as a mapping from $(D[0,1], \|\cdot\|_\infty)$ to $(D_\rho, \|\cdot\|_{\infty,\rho})$. The general strategy will be the same as in the previous subsection where we considered integral norms on the range space.

We first note that U_μ, which can be written as

$$(U_\mu f)(x+1) = \int_{(0, 1 \wedge x]} f\left(\frac{x}{y}\right) \mu(dy) \qquad \text{for all} \quad x \geqslant 0,$$

maps cadlag functions onto cadlag functions. To see this note that we can apply dominated convergence since

$$\left| f\left(\frac{x}{y}\right)\right| \leqslant \frac{\|f\|_{\infty,\rho}}{\rho(x/y)} \leqslant \frac{\|f\|_{\infty,\rho}}{\rho(1)} \qquad \text{if} \quad y \leqslant x,$$

as we assume that ρ is increasing. Also, the continuity of ρ implies that

$$\|f\|_{\infty,\rho} = \sup_{x \geqslant 1} \rho(x) |f(x-)|$$

for all cadlag functions f with $f(1-) = 0$, in particular for tail functions of signed measures on $[1, \infty)$ with total mass 0.

We have the following analogue of Lemma 3. Note that

$$c_\rho := \sup_{x \geqslant 0} \frac{\rho(1+x)}{\rho(1 \vee x)}$$

is finite by assumption (1).

LEMMA 9. (i) *For all* $f \in D_\rho$,

$$\|U_\mu f\|_{\infty,\rho} \leqslant C(\mu,\rho)\,\|f\|_{\infty,\rho}.$$

(ii) *For all* $f \in D[0,1]$ *and all nonnegative measures* ν *concentrated on* $[1,\infty)$,

$$\|T(f,\nu)\|_{\infty,\rho} \leqslant c_\rho\,\|f\|_\infty\,\|\nu\|_{\infty,\rho}.$$

Proof. The first part is immediate from

$$\begin{aligned}\|U_\mu f\|_{\infty,\rho} &\leqslant \sup_{x\geqslant 0}\,\rho(1+x)\int_{(0,\,1\,\wedge\,x]}\frac{1}{\rho(x/y)}\,\rho(x/y)\left|f\left(\frac{x}{y}\right)\right|\mu(dy)\\ &\leqslant \|f\|_{\infty,\rho}\sup_{x\geqslant 0}\,\rho(1+x)\int_{(0,\,1\,\wedge\,x]}\frac{1}{\rho(x/y)}\,\mu(dy),\end{aligned}$$

and the second part follows from

$$\begin{aligned}\|T(f,\nu)\|_{\infty,\rho} &\leqslant \sup_{x\geqslant 0}\,\rho(1+x)\int_{[x\,\vee\,1,\,\infty)}\left|f\left(\frac{x}{y}\right)\right|\nu(dy)\\ &\leqslant \|f\|_\infty\sup_{x\geqslant 0}\frac{\rho(1+x)}{\rho(1\vee x)}\,\sup_{x\geqslant 1}\rho(x)\,\nu([x,\infty))\\ &\leqslant \|f\|_\infty\,c_\rho\,\|\nu\|_{\infty,\rho}. \quad \blacksquare\end{aligned}$$

If $C(\mu,\rho)<1$ then V_μ can be defined as in Section 3.2, and Lemma 4 carries over without change, once $L_{1,\sigma}$ is replaced by D_ρ, and $e^\sigma m_1(\mu)<1$ by $C(\mu,\rho)<1$. We also need an analogue of Lemma 7.

LEMMA 10. *If* $\{\nu_\varepsilon\}_{0\leqslant\varepsilon\leqslant 1}$ *is a family of probability measures in* D_ρ *with*

$$\lim_{\varepsilon\to 0}\|\nu_\varepsilon-\nu_0\|_{\infty,\rho}=0,$$

then, for any $f\in D[0,1]$ *with* $f(1)=0$,

$$\lim_{\varepsilon\to 0}\|T(f,\nu_\varepsilon)-T(f,\nu_0)\|_{\infty,\rho}=0.$$

Proof. From

$$\left|\int_{[1,\infty)} 1_{[a,b)}\left(\frac{x-1}{y}\right) \nu_\varepsilon(dy) - \int_{[1,\infty)} 1_{[a,b)}\left(\frac{x-1}{y}\right) \nu_0(dy)\right|$$
$$\leqslant \left|\nu_\varepsilon\left(\left(\frac{x-1}{b}, \infty\right)\right) - \nu_0\left(\left(\frac{x-1}{b}, \infty\right)\right)\right|$$
$$+ \left|\nu_\varepsilon\left(\left(\frac{x-1}{a}, \infty\right)\right) - \nu_0\left(\left(\frac{x-1}{a}, \infty\right)\right)\right|$$

it follows that

$$\|T(1_{[a,b)}, \nu_\varepsilon) - T(1_{[a,b)}, \nu_0)\|_{\infty,\rho}$$
$$\leqslant \sup_{x\geqslant 1} \rho(x) \left|\nu_\varepsilon\left(\left(\frac{x-1}{b}, \infty\right)\right) - \nu_0\left(\left(\frac{x-1}{b}, \infty\right)\right)\right|$$
$$+ \sup_{x\geqslant 1} \rho(x) \left|\nu_\varepsilon\left(\left(\frac{x-1}{a}, \infty\right)\right) - \nu_0\left(\left(\frac{x-1}{a}, \infty\right)\right)\right|.$$

From the assumptions on ρ we obtain

$$\sup_{x\geqslant 1} \frac{\rho(x)}{\rho\left(\frac{x-1}{a} \vee 1\right)} < \infty$$

for all a in $(0, 1]$, so that the statement of the lemma holds for indicator functions of intervals $[a, b)$, $0 \leqslant a < b \leqslant 1$. We can now proceed as in the proof of Lemma 7. ∎

Let $\{\mu_\varepsilon\}_{0\leqslant\varepsilon\leqslant 1}$ again be a family of probability measures on $[0, 1]$ with $\mu_0(\{1\}) \neq 1$. We assume that $\rho(x) = O(e^{\sigma x})$ for some $\sigma < -\log m_1(\mu_0)$.

PROPOSITION 11. *If*

$$\lim_{\varepsilon\to 0} \sup_{0\leqslant x\leqslant 1} \left|\frac{1}{\varepsilon}(\mu_\varepsilon - \mu_0)((x, 1]) - g(x)\right| = 0$$

for some $g \in D[0, 1]$, *then*

$$\frac{1}{\varepsilon}(\Psi(\mu_\varepsilon) - \Psi(\mu_0)) \to V_{\mu_0} T(g, \Psi(\mu_0)) \qquad \text{in } D_\rho.$$

Proof. It follows from the proof of Lemma 5 and Markov's inequality that for all $\sigma_0 < -\log m_1(\mu_0)$ there exists an $\varepsilon_0 > 0$ and a $\kappa_0 < \infty$ such that for all $\varepsilon \in (0, \varepsilon_0)$

$$\Psi(\mu_\varepsilon)([x, \infty)) \leqslant \kappa_0 e^{-\sigma_0 x} \qquad \text{for all} \quad x \geqslant 1.$$

In particular, from our assumptions on ρ,

$$\limsup_{\varepsilon \to 0} \| \Psi(\mu_\varepsilon) \|_{\infty, \rho} < \infty.$$

As in the proof of Lemma 6, this implies the continuity of Ψ, and the statement of the proposition now follows on using exactly the same arguments as in the proof of Proposition 8. ∎

3.4. We now explain the step from the differentiability properties proved in the previous two subsections to the properties (AN) and (BW). The arguments are essentially the same as in Grübel and Pitts (1993), where a renewal theoretic setup is considered. Generally, the relationship between the differentiability of a functional and the transfer of asymptotic normality and the asymptotic validity of bootstrap confidence regions has been investigated by a number of authors, in particular Bickel and Freedman (1981), Gill (1989), Arcones and Giné (1992) and van der Vaart and Wellner (1996). We therefore content ourselves with a somewhat informal discussion; full details can easily be given along the lines of Grübel and Pitts (1993).

Let $(X_i)_{i \in \mathbb{N}}$ be a sequence of independent random variables with distribution μ, where $\mu([0, 1]) = 1$ and $\mu(\{1\}) \neq 1$. Let $\hat{F}_n$ be the distribution function associated with the empirical distribution $\hat{\mu}_n = \mu_n(X_1, ..., X_n)$. We then have

$$\sqrt{n}\, (\hat{F}_n - F) \to_d B \circ F, \tag{13}$$

where $B \circ F$ is a time-changed Brownian bridge. By the Skorohod–Wichura–Dudley construction (see, e.g., Shorack and Wellner, 1986, p. 47) we can obtain a pathwise version, i.e., on a suitable probability space we have

$$\sqrt{n}\, (F_n^\circ - F)(\omega^\circ) \to B^\circ \circ F(\omega^\circ) \qquad \text{in} \quad D[0, 1] \tag{14}$$

for all ω°, where the left- and right-hand side of (13) and (14), respectively, are equal to each other in distribution.

We now apply the differentiability of Ψ proved in Sections 3.2 and 3.3, respectively. The true distribution μ of the X-variables takes over the role of μ_0, $\varepsilon_n = n^{-1/2}$ and μ_{ε_n} corresponds to μ_n in the current notation. The differentiability results now yield, with respect to the appropriate norm and pointwise in each ω°,

$$\sqrt{n}\,(\Psi(F_n^\circ) - \Psi(F)) \to \Psi'_\mu(B^\circ \circ F).$$

Due to $\Psi(F_n^\circ) =_d \Psi(\hat{F}_n)$ $(= \hat{G}_n,)$ and $\Psi'_\mu(B^\circ \circ F) =_d \Psi'_\mu(B \circ F)$ this is the required weak convergence result for the empirical perpetuity.

For the analysis of the bootstrap we enlarge the above construction as in Bickel and Freedman (1981) and Gill (1989), the additional part modelling the resampling. For this, let $\{\xi_{ni} : n \in \mathbb{N}, 1 \leqslant i \leqslant n\}$ be an array of row-wise independent random variables, all uniformly distributed on the unit interval, and let $B^\dagger$ be a Brownian bridge, all defined on yet another probability space, such that

$$\sqrt{n}\,(H_n^\dagger - H)(\omega^\dagger) \to B^\dagger(\omega^\dagger) \qquad \text{in} \quad D[0, 1]$$

for all $\omega^\dagger$. Here $H_n^\dagger$ denotes the empirical distribution function associated with the nth row of the ξ-array and $H(t) = t$ for $0 \leqslant t \leqslant 1$. Then the random function $t \to H_n^\dagger(F_n^\circ(t, \omega^\circ))$ is uniformly distributed on the n^n distribution functions associated with the probabilities $\mu_n(X_{i_1}(\omega^\circ), ..., X_{i_n}(\omega^\circ))$, $(i_1, ..., i_n) \in I_n$, which implies that

$$R_n^\circ(\omega^\circ)(z) := P^\dagger(\sqrt{n}\,\|\Psi(H_n^\dagger(F_n^\circ(\cdot, \omega^\circ))) - \Psi(F_n^\circ(\cdot, \omega^\circ))\| \leqslant z),$$

regarded as a random function on Ω°, has the same distribution as the bootstrap estimator $\hat{R}_n$. It remains to show that R_n° tends to R P°-almost surely.

For this we again use the differentiability of Ψ. We have

$$\begin{aligned}\sqrt{n}\,(\Psi(H_n^\dagger(F_n^\circ)) - \Psi(F_n^\circ)) &= \sqrt{n}\,(\Psi(H_n^\dagger(F_n^\circ)) - \Psi(F)) \\ &\quad - \sqrt{n}\,(\Psi(F_n^\circ) - \Psi(F)). \end{aligned} \tag{15}$$

For the second term we obtain the limit $\Psi'_\mu(B^\circ \circ F))$. For the first term we use

$$\sqrt{n}\,(H_n^\dagger(F_n^\circ) - F) = \sqrt{n}\,(H_n^\dagger(F_n^\circ) - F_n^\circ) + \sqrt{n}\,(F_n^\circ - F). \tag{16}$$

A separate argument shows that the time change in the first term on the right-hand side does not matter, in the sense that we obtain the limit $B^\dagger \circ F$. Hence the right-hand side of (16) converges to $B^\dagger \circ F + B^\circ \circ F$. Using the differentiability of Ψ once more we therefore obtain

$$\sqrt{n}\,(\Psi(H_n^\dagger(F_n^\circ)) - \Psi(F)) \to \Psi'_\mu(B^\dagger \circ F + B^\circ \circ F).$$

Here, now, is the decisive step: Due to the linearity of the derivative a cancellation occurs, and we obtain the limit $\Psi'_\mu(B^\dagger \circ F)$ for the right-hand side of (15). Note that this no longer depends on ω°, that this quantity has the same distribution as $\Psi'_\mu(B^\circ \circ F)$, and that $\|\Psi'_\mu(B \circ F)\|$ has distribution function R.

3.5. Recall that $\rho(x) = e^{\sigma x}$ with some $\sigma < \log 2$. We need some more notation. For a probability measure μ on $[0, 1]$ let $\phi(\mu, \rho, \cdot): [0, \infty) \to \mathbb{R}$ be defined by

$$\phi(\mu, \rho, x) := \rho(1+x) \int_{(0, 1 \wedge x]} \frac{1}{\rho(x/y)} \mu(dy).$$

Clearly, $C(\mu, \rho) = \sup_{x \geqslant 0} \phi(\mu, \rho, x)$. Further, let unif$(a, b)$ be the uniform distribution on the interval (a, b). Finally, we introduce the function

$$\chi: [0, \infty) \to \mathbb{R}, \qquad \chi(\sigma) := e^{2\sigma} \int_0^1 e^{-\sigma/y}\, dy.$$

We claim that $\chi(\sigma) < 1$ for all $\sigma \in (0, \log 2)$, which is equivalent to

$$\int_\sigma^\infty \frac{1}{x^2} e^{-x}\, dx < \frac{1}{\sigma} e^{-2\sigma} \qquad \text{for} \quad 0 < \sigma < \log 2. \tag{17}$$

The Cauchy–Schwarz inequality yields

$$\begin{aligned} \left(\int_\sigma^\infty \frac{1}{x^2} e^{-x}\, dx \right)^2 &\leqslant \int_\sigma^\infty \frac{1}{x^4}\, dx \int_\sigma^\infty e^{-2x}\, dx \\ &= \frac{1}{6\sigma^3} e^{-2\sigma}, \end{aligned}$$

which shows that the inequality in (17) holds for $\sigma = \log 2$. From this (17) will follow if we can show that nowhere on the interval of interest is the derivative of the left-hand side of the inequality smaller than the derivative of the right-hand side, i.e., that

$$-\frac{1}{\sigma^2} e^{-\sigma} \geqslant -\frac{1}{\sigma^2} e^{-2\sigma} - \frac{2}{\sigma} e^{-2\sigma} \qquad \text{for} \quad 0 < \sigma < \log 2.$$

This in turn is equivalent to $e^\sigma \leqslant 1 + 2\sigma$ for $0 < \sigma < \log 2$, which is obviously true.

With $\mu = \text{unif}(0, \theta)$, $0 < \theta \leqslant 1$, we obtain

$$\begin{aligned}\sup_{x \geqslant \theta} \phi(\mu, \rho, x) &= \sup_{x \geqslant \theta} e^{\sigma} \frac{1}{\theta} \int_0^\theta e^{\sigma x(1 - 1/y)}\, dy \\ &= e^{\sigma(1+\theta)} \frac{1}{\theta} \int_0^\theta e^{-\sigma\theta/y}\, dy \\ &\leqslant e^{2\sigma} \int_0^1 e^{-\sigma/y}\, dy = \chi(\sigma),\end{aligned}$$

and for $0 < x < \theta$,

$$\begin{aligned}\phi(\mu, \rho, x) &= e^{\sigma(1+x)} \frac{1}{\theta} \int_0^x e^{-\sigma x/y}\, dy \\ &= e^{\sigma(1+x)} \frac{x}{\theta} \int_0^1 e^{-\sigma/y}\, dy \leqslant \chi(\sigma).\end{aligned}$$

In summary,

$$C(\text{unif}(0, \theta), \rho) \leqslant \chi(\sigma) \qquad \text{for all} \quad 0 < \theta \leqslant 1.$$

Obviously, for any given γ, the set of all μ satisfying $C(\mu, \rho) \leqslant \gamma$ is convex, and any continuous distribution which can be written as the weak limit of elements of this set, is also an element of this set. These closure properties can be used to lift the above statement to all probability measures on $[0, 1]$ with a decreasing density. Note that $\log 2$ is the upper bound for $-\log m_1(\mu)$ as μ ranges over the set of probability measures with decreasing density and support in $[0, 1]$.

REFERENCES

M. Aebi, P. Embrechts, and T. Mikosch, Stochastic discounting, aggregate claims, and the bootstrap, *Adv. Appl. Probab.* **26** (1994), 183–206.

M. A. Arcones and E. Giné, On the bootstrap of M-estimators and other statistical functionals, *in* “Exploring the Limits of Bootstrap” (R. LePage and L. Billard, Eds.), pp. 13–47, Wiley, New York, 1992.

P. Bickel and D. Freedman, Some asymptotic theory for the bootstrap, *Ann. Statist.* **9** (1981), 1196–1217.

N. G. de Bruijn, The asymptotic behaviour of a function occurring in the theory of primes, *J. Indian Math. Soc.* **15** (1951), 25–32.

D. Dufresne, The distribution of a perpetuity, with applications to risk theory and pension funding, *Scand. Actuarial J.* (1990), 39–79.

R. D. Gill, Non- and semi-parametric maximum likelihood estimators and the von Mises method (Part I), *Scand. J. Statist.* **16** (1989), 97–128.

C. M. Goldie and R. Grübel, Perpetuities with thin tails, *Adv. Appl. Probab.* **28** (1996), 463–480.

R. Grübel, Hoare's selection algorithm: A Markov chain approach, *J. Appl. Probab.* **35** (1998), 36–45.

R. Grübel and S. M. Pitts, Nonparametric estimation in renewal theory I: The empirical renewal function, *Ann. Statist.* **21** (1993), 1431–1451.

R. Grübel and U. Rösler, Asymptotic distribution theory for Hoare's selection algorithm, *Adv. Appl. Probab.* **28** (1996), 252–269.

H. Heuser, "Functional Analysis," Wiley, Chichester, 1982.

S. M. Pitts, Nonparametric estimation of the stationary waiting time distribution function for the GI/G/1 queue, *Ann. Statist.* **22** (1994a), 1428–1446.

S. M. Pitts, Nonparametric estimation of compound distributions with applications in insurance, *Ann. Inst. Statist. Math.* **46** (1994b), 537–555.

G. R. Shorack and J. A. Wellner, "Empirical Processes with Applications to Statistics," Wiley, New York, 1986.

A. W. van der Vaart and J. A. Wellner, "Weak Convergence and Empirical Processes," Springer-Verlag, New York, 1996.

W. Vervaat, On a stochastic difference equation and a representation of non-negative infinitely divisible random variables, *Adv. Appl. Probab.* **11** (1979), 750–783.

Printed in Belgium